Systematic Entomology
DOI: 10.1111/syen.12397
Accepted 15 September 2019
Final draft version/ Author's Accepted Manuscript

Genomic evidence suggests *Mesapamea remmi* is an imaginary species (Lepidoptera: Noctuidae)

Pasi Sihvonen (corresponding author), Finnish Museum of Natural History LUOMUS, University of Helsinki, Helsinki, Finland, pasi.sihvonen@helsinki.fi
Kyung Min Lee, Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland
Karl-Erik Lundsten, Rintamamiehentie 7, Espoo, Finland
Marko Mutanen, Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland

Abstract. Taxonomic status of the noctuid moth *Mesapamea remmi* Rezbanyai-Reser (Lepidoptera) has remained controversial since its description in 1985. Based on morphology, it has been considered a valid species or a hybrid between *Mesapamea secalis* (Linnaeus) and *Mesapamea didyma* (Esper). We studied this case of uncertain identity by using traditional Sanger sequencing techniques (COI and seven nuclear genes) as well as double digest restriction-site associated DNA (ddRAD) sequencing, which resulted in analyses including 9402 loci and 1 512 976 bp across the genome. Our analyses showed that genomic data do not support the hybrid hypothesis and that *M. remmi* is not separated genetically from *M. secalis*; both are clearly distinct from *M. didyma*. Reproductive organs of *M. remmi* males are unique and diagnosable, whereas females are teratological in the sense that ductus seminalis is missing and corpus bursae is malformed, the latter being connected directly to the ovipore. These data support the view that *M. remmi* is not a valid species and we consider *M. remmi* Rezbanyai-Reser, 1985 to be a junior synonym of *M. secalis* (Linnaeus, 1758). We demonstrate that genomic approaches provide an efficient way to address various difficult and long-standing taxonomic issues, in this case oversplitting of species.

Introduction

Species as a taxonomic concept have a particular role in both biological research and in many areas of human society, such as legislation, agriculture and food industry. Therefore, analytical tools that enhance our ability to identify species are of great importance. During the last decade, progress in sequencing technology has been unprecedentedly rapid so that it is now possible both to recover genomic-scale data from specimens (for review on Lepidoptera, see Triant *et al.*, 2018) and to identify thousands of specimens by their genetic fingerprints (DNA barcodes) in a straightforward manner and short time period (Hebert *et al.*, 2003; Janzen & Hallwachs, 2016). Broad DNA barcoding campaigns have revealed putative cryptic species in many groups (e.g. Mutanen *et al.*, 2013; Huemer & Karsholt, 2018) but also possible cases of taxonomic oversplitting of species (Mutanen *et al.*, 2016). Few studies have focused on such taxonomic uncertainties by using genomic means, despite the availability of various efficient methods.

Western Palearctic *Mesapamea secalis* (Linnaeus) (Lepidoptera: Noctuidae) is an occasional yet important pest on rye (*Secale cereale* Linnaeus), whose larvae destroy seedlings in the autumn, and continue eating the uppermost leaf sheath of older plants the following spring, causing the diagnostic 'white head' appearance.Damage can be locally significant. For instance, historical records from Finland include damage of 50% or even 80–100% (Vappula, 1962 and references therein), and in Great Britain it is considered as one of the most important pests of cultivated pasture (Carter, 1984). *Mesapamea secalis* can cause a similar 'white head' appearance on wheat (*Triticum aestivum* Linnaeus), *Phleum pratense* Linnaeus and

Alopecurus Linnaeus (Mikkola & Jalas, 1979), and it also feeds on oat (*Avena* Linnaeus) and maize (*Zea* Linnaeus) (Carter, 1984).

Estonian Hans Remm discovered that *M. secalis* is a cryptic species pair, and he described a new species called *M. secalella* Remm, which is currently considered a junior synonym of *M. didyma* (Esper) (nomenclature used in our article follows Karsholt & Stadel Nielsen, 2013). These species cannot be identified based on external features (Remm, 1983). After the initial finding, these two species have been researched intensively, partly because of the damage they cause on rye and wheat, but also to understand their distributions and relative abundances (Zilli *et al.*, 2005 and references therein). One such study led to the discovery of an additional taxon in this complex from Switzerland: *M. remmi* Rezbanyai-Reser. Although *M. remmi* is morphologically distinct (Rezbanyai-Reser, 1985; see Figs 1–6), its taxonomic status has remained unclear. Recently, *M. remmi* has been considered a valid species (van Nieukerken & Karsholt, 2013) or, predominantly, as an occasional hybrid between *M. secalis* × *M. didyma* on the grounds that: (i) some characters of *M. remmi* are intermediate between the two parent species (particularly the clavus in the male genitalia), while others are unique (e.g. swollen V structure of the lamella antevaginalis in the female genitalia); (ii) females may be teratological, with the ductus bursae and the corpus bursae being absent; (iii) *M. remmi* has so far been found only in areas where both assumed parent species occur; and (iv) *M. remmi* specimens are extremely rare (ratio *secalis*:*didyma*:*remmi* =206:124:1) (Zilli *et al.*, 2005; Hausmann *et al.*, 2011; Huemer, 2013; Gaedike *et al.*, 2017; Huemer *et al.*, 2019, abundance data from Rezbanyai-Reser, 1985). Much of the evidence is circumstantial and the hybrid hypothesis has not been tested by breeding and backcrossing experiments.

We first tried to study the taxonomic problem by using Sanger-based sequencing of the mitochondrial COI gene and by seven nuclear genes – cytochrome *c* oxidase subunit I (COI), Elongation factor 1-alpha 1 (EF-1a), glyceraldehyde 3-phosphate dehydrogenase (GAPDH), isocitrate dehydrogenase (IDH), malate dehydrogenase (MDH), ribosomal protein S5 (RpS5), wingless and Dopa decarboxylase (DDC) (Table 1) – but these did not provide sufficient information to assess the taxonomic status of *M. remmi* (see the Results section). Thus, we performed a study using double digest RAD (ddRAD) sequencing, an approach that enables rapid recovery of thousands of orthologous loci from specimens (Peterson *et al.*, 2012). The following possible scenarios were addressed: (i) *M. remmi* is a distinct species; (ii) *M. remmi* is hybrid between *M. secalis* and *M. didyma*; and (iii) *M. remmi* is synonymous with *M. secalis* or *M. didyma*. We hypothesized observing the following patterns under each scenario, respectively: (i) unique nucleotide substitution to occur in each taxa or at least their separation at genomic level; (ii) *M. remmi* to be genetically intermediate between the species and heterozygous at sites that show fixed alleles between *M. secalis* and *M. didyma*; and (iii) *M. remmi* to be indistinguishable from one of the other species at the genomic level.

Materials and methods

Data for the study were obtained from southern Finland, where *M. secalis*, *M. didyma* and *M. remmi* have been reported to co-occur (e.g. Sihvonen, 2000; Lundsten & Sihvonen, 2003). A collecting campaign was carried out in 2007, when all *Mesapamea* specimens (3475 specimens) caught in 23 automatic traps (nine light traps, 14 bait traps) were morphologically identified based on the reproductive organs (Rezbanyai-Reser, 1985; Zilli *et al.*, 2005). Light traps were timer-operated 160W mixed-light 'Jalas type' (Jalas, 1992), placed 0.5–1 m above ground, and emptied about once a week. Bait traps with red wine baits were placed in trees 3–5 m above ground, and emptied and rebaited about once a week. These data were used to study the abundance of each taxon. These observations were supplemented by individual specimens of each species from other years, constituting the material used in genetic and morphological analyses. In the ddRAD analyses, both sexes were available for *M. secalis* and *M. didyma*, but only females (*n* =6) were included for *M. remmi* (Appendix S2).

Molecular analyses and bioinformatics

DNA extractions were performed using the DNeasy Blood & Tissue Kit (Qiagen, Germany) according to the manufacturer's instruction. The quantity of gDNA extracts was checked using the Quant-iT™ PicoGreen® dsDNA assay Kit (Molecular Probes, Eugene, OR, U.S.A.). All PCR and sequencing protocols followed Wahlberg &Wheat (2008), except for PCR purification which was carried out with ExoSAP-IT (Affymetrix, Santa Clara, CA, USA) and Sephadex columns (Sigma-Aldrich, St. Louis, MO, U.S.A.) and sequencing that was performed with an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA, U.S.A.). All sequences for each taxon were aligned and edited using bioedit 7 (Hall, 1999). All sequences are available at the NCBI GenBank (accession numbers are provided in Appendix S3).

The ddRAD library construction followed protocols described in Lee et al. (2018) with two exceptions: gDNA was digested with PstI and MspI, and the size distribution and concentration of the pools were measured with a Bioanalyzer (Agilent Technologies). Raw demultiplexed reads are available in the NCBI Sequence Read Archive (SRA) under the Bioproject PRJNA505223.

Rawpaired-end reads were demultiplexedwith no mismatches tolerated using their unique barcode and adapter sequences with ipyrad v.0.7.23 (Eaton & Overcast, 2016). All ipyrad defaults were used, with the following exceptions: the minimum depth at which majority-rule base calls are made was set to 3, the cluster thresholds were set to 0.80, 0.85, 0.90 and 0.95, and the minimum numbers of samples that must have data at a given locus for it to be retained were set to 4, 10, and 17.

Data analyses

Sequences of COI and the seven nuclear genes were examined for fixed substitutions between species in sequence alignments using mega 7.0 (Kumar *et al.*, 2016).

Maximum likelihood (ML) trees were constructed for the concatenated ddRAD data. Maximum likelihood trees were inferred in raxml v.8.2.0 (Stamatakis, 2014) with bootstrap support estimated by 1000 replicates of rapid bootstrap (Stamatakis *et al.*, 2008) from the unpartitioned super alignment under GTR+CAT model. We visualized the resulting ML tree with the best likelihood score and assessed bootstrap support using Figtree v.1.4.2 (Rambaut, 2015).

We then applied a number of filters to identify candidate diagnostic single nucleotide polymorphisms (SNPs) for detecting fixed allelic differences between the species.We focused on loci genotyped for all individuals assayed (0% missing data) in 17 specimens (both males and females) and on RAD tags containing only one biallelic SNP.We collected only SNPs fixedwithout appearance of any heterozygosity at the sites.

We used admixture analysis implemented in structure v.2.3.1 (Pritchard *et al.*, 2000) using SNP frequency data to investigate genomic variation between individuals. Ten replicates were run at each value of K. Each run had a burn-in of 10K generations followed by 20K generations of sampling. Replicates were permuted in the program clumpp (Jakobsson & Rosenberg, 2007) according to the ad hoc ΔK statistics (Evanno *et al.*, 2005). structure results were visualized using the program distruct (Rosenberg, 2004).

Reproductive organs of all taxa were prepared following standard methods (Hardwick, 1950). The vesica was everted via the opening of the ductus ejaculatorius (Lafontaine & Mikkola, 1998) or via caecumwhichwas cut open by placing the aedeagus inside a hypodermic syringe (Sihvonen, 2001). Comparative morphology was used to examine genitalia structures. Female corpus bursae was photographed in ethanol to show it as fully expanded, and it was transferred to Euparal afterwards. Taxa shown in the genitalia plates were photographed in the ventral view with a Leica DM1000 microscope

(Wetzlar, Germany) and an integrated Leica DF295 digital camera. Numerous structures were photographed in two to six images of different depth of focus and combined into single images using image-stacking as available in Adobe photoshop 19. The final plates were compiled in CorelDraw (v.19).

Data accessibility

Voucher specimens are stored in the private collection of K-EL, Espoo (these will be deposited in a public collection in due course), and in the Finnish Museum of Natural History, Helsinki, Finland. Specimens in K-EL's collection are available for loan via the Finnish Museum of Natural History or directly from the author.

Project data have been deposited at NCBI Sequence Read Archive under the accession PRJNA505223.

GenBank accession numbers for sequences used in this study are provided in Supplement 03.

The datasets supporting this article have been uploaded as part of the supplementary material.

Ethical statement

This work includes research on invertebrates, which do not need ethical permissions. Material was collected from Finland, outside protected areas, and the national legislation allows such material to be collected, stored and researched without permits. All collected material was studied in Finland, and therefore no import/export permits were needed. All data were collected before 12 October 2014 and therefore the genomic data are not governed by the ABS treaty of the Nagoya Protocol.

The studied *Mesapamea* are common (Table S1) and found in large numbers in Finland. Collection during the course of this work has not endangered the sampled populations, because samples represented a minute fraction of populations.

Results

The mitochondrial COI gene was sequenced for 1476 bp, of which four sites showed a fixed nucleotide substitution between *M. secalis* and *S. didyma*. Five additional sites were nearly fixed between *M. secalis*/*M. remmi* and *M. didyma*, but in each case with one deviating individual. No fixed substitutions were found between *M. secalis* and *M. remmi*. The seven nuclear genes (3674 bp in all) did not show any fixed substitutions between *M. secalis* and *M. remmi*, thus not permitting us to test whether *M. remmi* is heterozygous for such sites (Appendix S3). An overview of the Sanger-based sequencing and observed numbers of fixed substitutions between species are provided in Table 1.

For ddRAD data, we obtained 2.66 million reads per individual on average, of which 81.3% were retained after stringent quality-filtering steps (Table 2). After filtering and clustering at 90% sequence similarity, we recovered 9402 putative orthologous loci shared across more than four samples, for a total length of 1 512 976 bp (Appendix S4). These data include 27 539 SNPs, of which 11 060 are parsimony-informative sites (PIS).

Phylogenetic analysis using the concatenated RAD dataset of the most variable ('ddRAD_c90m4'), which includes the highest proportion of variable sites and ('ddRAD_c90m17'), which contains fewer SNPs and no missing data, produced robust support for the relationship between *M. didyma* and *M. remmi* + *M. secalis* (Fig. 7A; Appendix S5). In the ML trees, the two major clades correspond to *M. didyma* and *M. remmi* + *M. secalis*; both clades were supported by 100% bootstrap values.

Structure also identified two discrete clusters, which correspond to *M. didyma* and *M. remmi* + *M. secalis* (Fig. 7B, *K* =2). If three genetic clusters were assumed, *M. remmi* and *M. secalis* were completely admixed with each other (Fig. 7B, *K* =3).

A total of 334 putative RAD loci had exactly one biallelic putative SNP and were genotyped in all 17 individuals of *Mesapamea*. The data include a total of 894 SNPs, of which 479 are PIS. The SNPs occur at 2.68 SNPs per locus on average. Of these, we identified 10 fixed differences between *M. didyma* and *M. remmi* + *M. secalis*, which may prove to be candidate lineage-specific SNPs (Table 3).

Observations of *Mesapamea* species in random samples in southern Finland resulted in the following ratios for *M*. *secalis*:*M. didyma*:*M*. *remmi*: 3193 specimens (91.9%):262 specimens (7.5%):20 specimens (0.6%). *Mesapamea remmi* is extremely rare. Details are provided in Table 4.

The reproductive organs of *M. remmi* males (Fig. 1) and females (Fig. 4) are distinct and diagnosable, but females are teratological in the sense that the ductus seminalis is missing and the corpus bursae is malformed (often frill-shaped); the latter is connected directly to the ovipore (Fig. 4). As a result, females are incapable of reproducing.

Discussion

Of the three possible scenarios hypothesized in the introduction, scenario number (iii) was supported: *M. remmi* is synonymous with *M. secalis*. This is supported by mitochondrial COI, seven nuclear genes and by genomic data, where *M. remmi* is indistinguishable from *M*. *secalis*, and *M*. *remmi* and *M*. *secalis* specimens were admixed (Fig. 7B, *K* =3). The genomic data does not support the hybrid hypothesis because *M*. *remmi* is not genetically intermediate between *M. didyma* and *M*. *secalis* and it is not heterozygous at sites that show fixed alleles between *M. didyma* and *M*. *secalis*. These genomic data support the view that *M. remmi* is not a valid species and we consider *M. remmi* Rezbanyai-Reser to be a junior synonym of *M. secalis* (Linnaeus). This result has implications because we now have evidence supporting the view that *M. secalis* and *M. didyma* form a species pair with agricultural importance, not a pest complex of three species. Earlier literature on these taxa can now be examined critically and with this refined knowledge. Other researchers can now better investigate potential biological controls of these species and gain a clearer understanding of their distribution and their life histories.

With regard to the teratological structures in the female reproductive organs of taxon *M*. *remmi*, these are peculiar in the sense that all examined specimens (*n* =6) are similar: the ductus seminalis is missing between the bursa copulatrix and the genital chamber (*n* =6), indicating that it is not a biologically reproductive species. Zilli *et al.* (2005) report that the corpus bursae of taxon *remmi* is often absent, and they illustrate such a specimen.We did not observe an absence of the corpus bursae in the material studied.

Our genomic analyses cannot confirm why the male and female genitalia of taxon *remmi* are morphologically distinct and diagnosable from *M. secalis* and *M. didyma* (Figs 1–6). Potential explanations could be genitalia polymorphism, chromosomal disorder(s) or other hereditary disorder(s) with a frequency of <1%. The reported cases of genitalia polymorphism in Lepidoptera are rare, and it may be partially a result of circular reasoning as species are often delimited by differences in genital characteristics (Hausmann, 1999; Mutanen&Kaitala, 2006). An understanding of potential chromosomal or other hereditary disorders, which are observable in the genitalia phenotype, would require karyological analysis in *Mesapamea*. In the Noctuidae, the haploid number *n* =31 is considered the modal chromosome number (Werner, 1975; de Prins & Saitoh, 2003).

With respect to the relative abundance of *M. remmi*, our data report 0.6% specimens as belonging to this taxon (Table 4), which is similar to what was observed in earlier studies [0.3% (Rezbanyai-Reser, 1985) and 0.3% (Rezbanyai-Reser, 1989)].

Taxonomy and species concepts play a central role in our understanding of the world's described biodiversity and two opposite processes of taxonomy are fundamental: species descriptions and synonymizations. Both of these may be inaccurate, the former including oversplitting or undersplitting of species. Oversplitting is apparently not a rare phenomenon, at least in areas with a long taxonomic tradition such as Europe. Recently, Mutanen *et al.* (2016), using an extensive DNA barcode dataset, reported nearly 150 possible cases where conspecificity may be involved, equalling 31.8% of all nonmonophyletic cases identified in that study. Taxonomists have had limited tools to tackle such difficult issues, but recent advances in genomics, such as the case presented in our study, are a welcome resource and allow more detailed studies and conclusions to be made.

Mesapamea insolita Rezbanyai-Reser is yet another controversial taxon in the studied species complex (Rezbanyai-Reser, 1996). Like *M. remmi*, it was also described from Switzerland and subsequent authors have considered it a hybrid, and formally it has been treated as a junior synonym of *M. secalella* (Zilli *et al.*, 2005) [= *didyma*, according to Karsholt & Stadel Nielsen's classification (Karsholt & Stadel Nielsen, 2013)] or a deformed, teratological specimen. Only the holotype male is known so far and it has not been subject to genetic analyses. An approach similar to our study may help to solve this and many other taxonomic problems.

Authors' contributions

PS conceived, designed and coordinated the study, carried out morphological work and drafted the manuscript. KML carried out the molecular lab work, performed the bioinformatics, data analysis and drafted the manuscript. K-EL collected the material, identified the specimens, summarized adundance data and drafted the manuscript. MM conceived, designed and coordinated the study, participated in data analysis and drafted the manuscript. All authors gave final approval for publication.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Metadata of the specimens illustrated in Figs 1–6.
Appendix S2. Metadata of the specimens used in genomic studies.
Appendix S3. Sequencing success in studied taxa for mitochondrial COI and seven nuclear genes.
Appendix S4. Sequence information in the ddRAD and mitochondrial COI data matrices.
Appendix S5. Maximum likelihood unrooted tree based on the most conservative data matrix ('ddRAD_c90m17').

Acknowledgements

We thank Laura Törmälä for the efficient work in the lab and CSC – IT Centre for Science, Finland, for computational resources and Victoria Twort (Finnish Museum of Natural History, University of Helsinki) for revising the English language. This work was supported by the Academy of Finland (grant number 277984 allowed toMMhttps://doi.org/10.13039/501100002341) and the Kvantum Institute (University of Oulu) to KML. The authors declare there are no conflicts of interest. All authors confirm that there are no disputes over the ownership of the data presented in the paper and all contributions have been attributed appropriately, via co-authorship or acknowledgement, as appropriate to the situation.

References

Carter, D.J. (1984) *Pest Lepidoptera of Europe with Special Reference to the British Isles*. W. Junk Publishers, Dordrecht. 431 p.

Eaton, D.A.R. & Overcast, I. (2016) *ipyrad: interactive assembly and analysis of RADseq data sets*. URL http://ipyrad.readthedocs.io/ [accessed on 1 March 2019].

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Gaedike, R., Nuss, M., Steiner, A. & Trusch, R. (2017) Verzeichnis der Schmetterlinge Deutschlands (Lepidoptera). 2. überarbeitete Auflage. *Entomologischen Nachrichten und Berichte*, **21**, 1–362.

Hall, T. (1999) BioEdit: a user-friendly biological sequences alignment editor and analysis program for windows 95/98/NT. *Nucleic Acid Symposium Series*, **41**, 95–98.

Hardwick, D.F. (1950) Preparation of slide mounts of Lepidopterous genitalia. *Canadian Entomologist*, **82**, 231–235.

Hausmann, A. (1999) Falsification of an entomological rule: polymorphic genitalia in geometrid moths. *Spixiana*, **22**, 83–90.

Hausmann, A., Haszprunar, G., Segerer, A., Speidel, W., Behounek, G. & Hebert, P.D.N. (2011) Now DNA-barcoded: the butterflies and larger moths of Germany (Lepidoptera: Rhopalocera, Macroheterocera). *Spixiana*, **34**, 47–58.

Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003) Biological identifications throughDNAbarcodes. *Proceeedings of the Royal Society London, Series B*, **270**, 313–321.

Huemer, P. (2013) *Die Schmetterlinge Österreichs (Lepidoptera)*. Systematische und faunistische Checkliste. Tiroler Landesmuseen-Betriebsgesellschaft m.b.H, Innsbruck. 304 p.

Huemer, P. & Karsholt, O. (2018) Revision of the genus *Megacraspedus* Zeller, 1839, a challenging taxonomic tightrope of specie delimitation (Lepidoptera, Gelechiidae). *ZooKeys*, **800**, 1–278. https://doi.org/10.3897/zookeys.800.26292.

Huemer, P.,Wieser, C., Stark,W., Hebert, P.D.N. &Wiesmair, B. (2019) DNA barcode library of megadiverse Austrian Noctuoidea (Lepidoptera) – a nearly perfect match of Linnean taxonomy. *Biodiversity Data Journal*, **7**, e37734.

Jakobsson, M. & Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jalas, I. (1992) Perhostenkeräilijän opas [Lepidopterist's guide]. 3rd ed. Otava, Keuruu. 272 p.

Janzen, D.H. & Hallwachs, W. (2016) DNA barcoding the Lepidoptera inventory of a large complex tropical conserved wildland, area de Conservacion Guanacaste, northwestern Costa Rica. *Genome*, **59**, 641–660.

Karsholt, O. & Stadel Nielsen, P. (2013) *Revideret fortegnelse over Danmarks Sommerfugle*. Lepidopterologisk Forening, København. 120 p.

Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0. *Molecular Biology and Evolution*, **33**, 1870–1874.

Lafontaine, J.D. & Mikkola, K. (1998) Lås-och-nyckel systemen i de inre genitalierna av Noctuidae (Lepidoptera) som taxonomiska kännetecken. *Entomologiske Meddelelser*, **55**, 161–167.

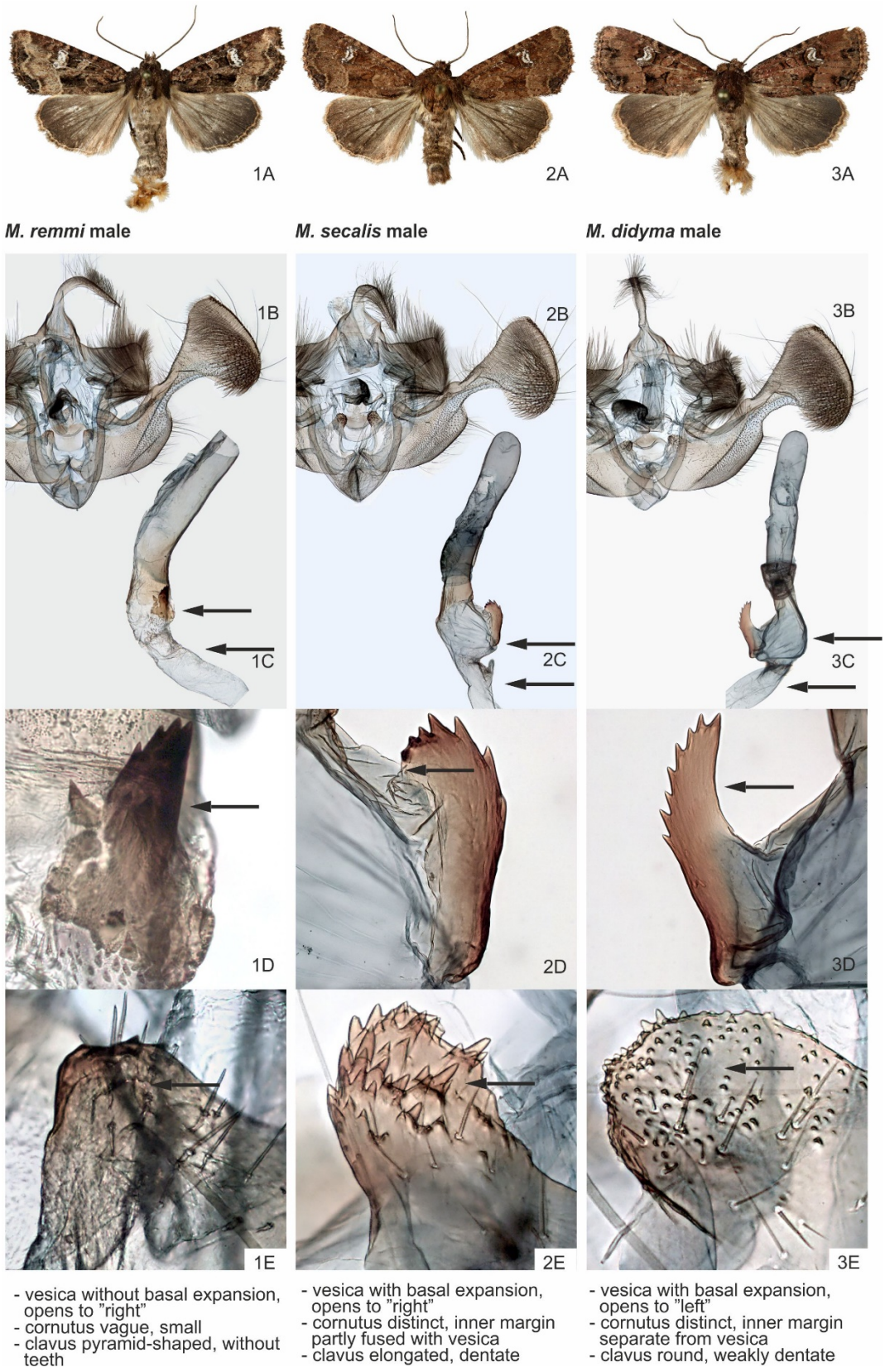
Lee, K.M., Kivelä, S.M., Ivanov, V., Hausmann, A., Kaila, L.,Wahlberg, N. & Mutanen, M. (2018) Information dropout patterns in restriction site associated DNA phylogenomics and a comparison with multilocus sanger data in a species-rich moth genus. *Systematic Biology*, **67**, 925–939.

Lundsten, K.-E. & Sihvonen, P. (2003) *Mesapamea remmi* Rezbanyai-Reser – Suomelle uusi yökköslaji [*Mesapamea remmi* Rezbanyai-Reser found in Finland] (Lepidoptera: Noctuidae). *Baptria*, **28**, 18–19. In Finnish with English abstract.

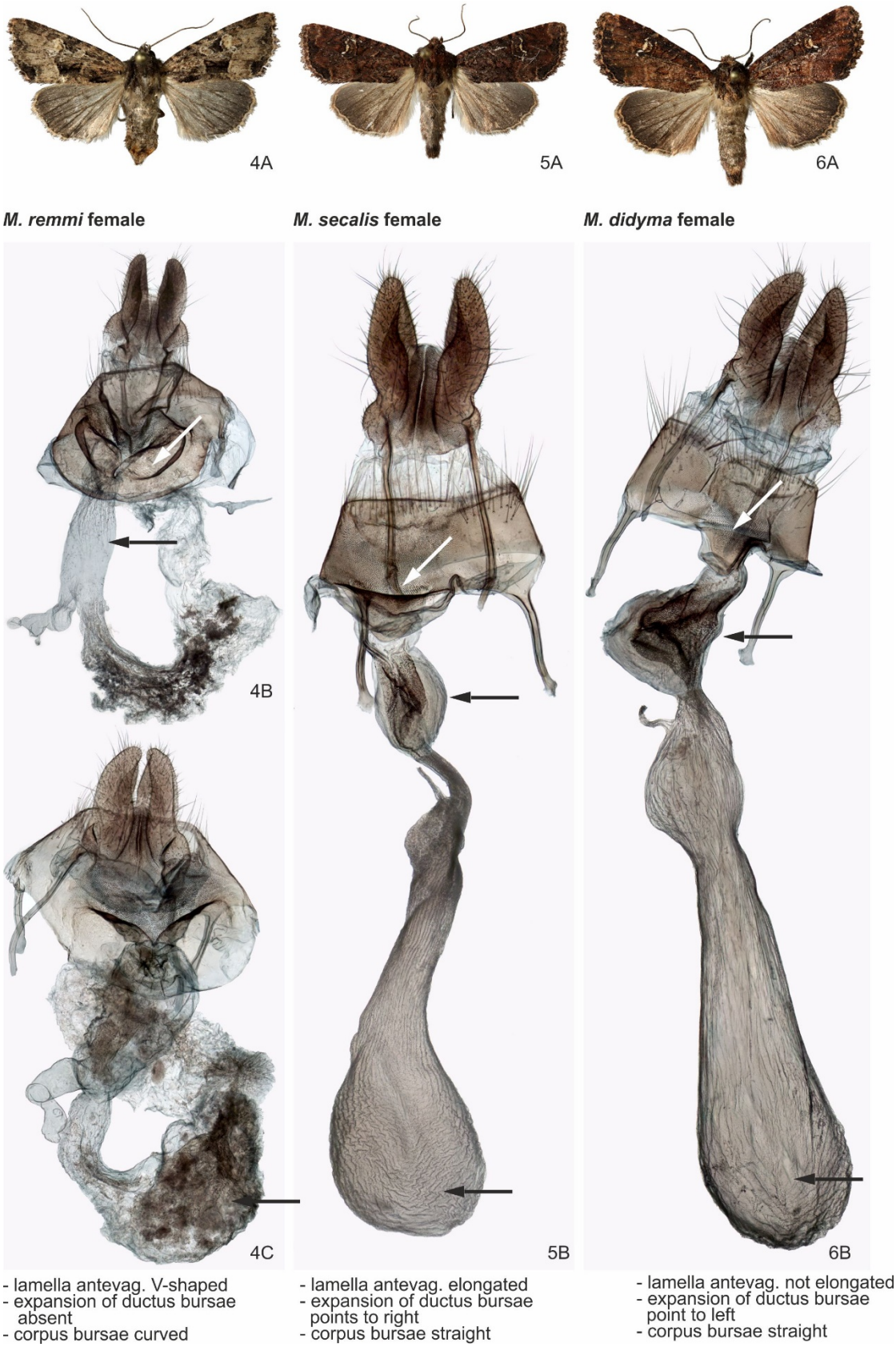
Mikkola, K. & Jalas, I. (1979) *Suomen perhoset, yökköset 2 (Finnish Lepidoptera, Noctuidae 2, in Finnish)*. Otava, Keuruu. 304 p.

Mutanen, M., Kaila, L. & Tabell, J. (2013) Wide-ranging barcoding aids discovery of one-third increase of species richness in presumably well-investigated moths. *Scientific Reports*, **3**, 1–7. https://doi.org/10.1038/srep02901.

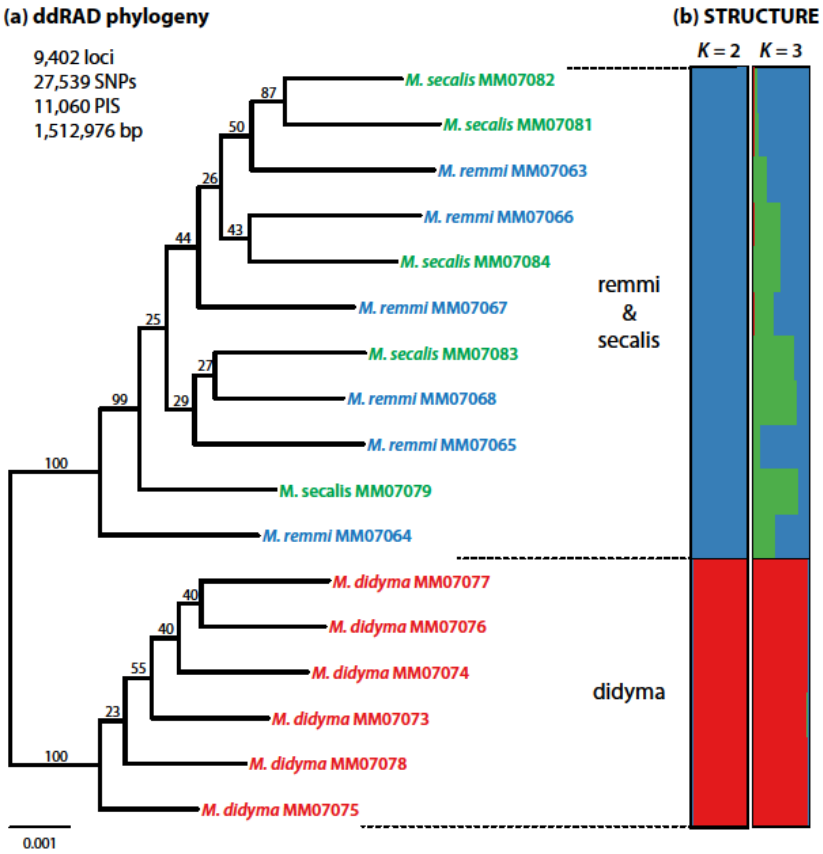
Mutanen, M. & Kaitala, A. (2006) Genitalia variation in a dimorphic moth *Selenia tetralunaria* (Lepidoptera: Geometridae). *Biological Journal of the Linnean Society*, **87**, 297–307.
Mutanen,M., Kivelä, S.M., Vos, R.A., Doorenweerd, C., Ratnasingham, S. *et al.* (2016) Species-level Para- and polyphyly in DNA barcode gene trees: strong operational bias in European Lepidoptera. *Systematic Biology*, **65**, 1024–1040.
van Nieukerken, E.J. & Karsholt, O. (2013) Lepidoptera, moths. *Fauna Europaea*, version 2017.06. URL https://fauna-eu.org [accessed 1 March 2019].
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PloS One*, **7**, e37135. https://doi.org/10.1371/journal.pone.0037135.
de Prins, J. & Saitoh, K. (2003) Karyology and sex determination. *Handbook of Zoology*, *Lepidoptera, moths and butterflies*, Vol. **2** (ed. by N.P. Kristensen), pp. 449–468. Walter de Gruyter, Berlin.
Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Rambaut, A. (2015) *FigTree*, v.1.4.2: tree figure drawing tool.Molecular evolution, phylogenetics and *Epidemiology* URL http://tree.bio.ed.ac.uk/software/figree/ [accessed1 February 2019].
Remm, H. (1983) New species of Noctuidae (Lepidoptera) from the USSR. *Entomologicheskoe Obozrenie*, **62**, 596–600.
Rezbanyai-Reser, L. (1985) *Mesapamea*-Studien II. *Mesapamea remmi* sp. n. aus der Schweiz, sowie Beiträge zur Kenntnis der westpalaearktischen Arten der Gattung *Mesapamea* Heinicke, 1959 (Lep., Noctuidae). *Entomologische Berichte Luzern*, **14**, 127–148.
Rezbanyai-Reser, L. (1989) *Mesapamea*-Studien III: Angaben zum Vorkommen, zur Häufigkeit und Phänologie von *M. secalis* L., *didyma* Esp. (= *secalella* Remm) und *remmi* Rezb.-Reser, aufgrund kontinuierlicher Lichtfallenfangergebnisse in der Schweiz von 1983-87 (Lepidoptera, Noctuidae). *Entomologische Berichte Luzern*, **21**, 67–103.
Rezbanyai-Rser, L. (1996) *Mesapamea*-Studien IX. *Mesapamea insolita* n. sp. aus der Südschweiz sowie die bisherigen Fundangaben von *Mesapamea remmi* Rezbanyai-Reser 1985 in Europa (Lepidoptera: Noctuidae). *Entomologische Zeitschrift (Essen)*, **106**, 81–92.
Rosenberg, N. (2004) DISTRUCT: a program for the graphic display of population structure. *Molecular Ecolology Notes*, **4**, 137–138.
Sihvonen, P. (2000) *Mesapamea*-tähkäyökköset Suomessa (Lepidoptera: Noctuidae) [*Mesapamea* noctuids in Finland] (Lepidoptera: Noctuidae). *Baptria*, **25**, 146–153. In Finnish with English abstract.
Sihvonen, P. (2001) Everted vesicae of the *Timandra griseata* group: methodology and differential features (Geometridae, Sterrhinae). *Nota Lepidopterologica*, **24**, 57–63.
Stamatakis, A. (2014) RaxML vesion 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.
Stamatakis, A., Hoover, P. & Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, **57**, 758–771.
Triant, D.B., Cinel, S.C. & Kawahara, A.Y. (2018) Lepidoptera genomes: current knowledge, *gaps and future directions. Current Opinion in Insect Science*, **25**, 99–105. https://doi.org/10.1016/j.cois.2017.12.004.
Vappula, N.A. (1962) Pests of cultivated plants in Finland. *Annales Agriculturae Fenniae* **1**, Supplementum 1 (English edition), Seria Animalia Nocentia 5. 239 p.
Wahlberg, N. & Wheat, C.W. (2008) Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Systematic Biology*, **57**, 231–242.
Werner, L. (1975) Zur evolution des Karyotyps bei den Lepidopteren. Die Chromosomenzahlen der Noctuiden. *Genetica*, **4**, 377–395.
Zilli, A., Ronkay, L. & Fibiger, M. (2005) *Noctuidae Europaeae*, Vol. **8**. Apameini. Entomological Press, Sorø. 323 p.



Figs 1–3. Diagnostic male genitalia characters (indicated) of *Mesapamea* species: 1, *M. remmi*; 2, *M. secalis* and 3, *M. didyma*. Aedeagi with everted vesicas are shown in the copulation position relative to the female genitalia (Figs 4–6), and diagnostic species-level characters are indicated and explained. Metadata of illustrated specimens are given in Appendix S1.



Figs 4–6. Diagnostic female genitalia characters (indicated) of *Mesapamea* species: 4, *M. remmi*; 5, *M. secalis*; and 6, *M. didyma*. Diagnostic species-level characters are indicated and explained. Metadata of illustrated specimens are given in Appendix S1.



Figs 7. Phylogenetic relationships and population structuring of *Mesapamea* based on the 'ddRAD_c90m4' data matrix, which consisted of 27 539 SNPs in 1 512 976 bp. (A) The maximum likelihood tree was inferred in raxml with 1000 bootstrap replicates. The bootstrap values are indicated above branches. (B) Results of the population structure analyses with posterior probability plots of individual assignments to the inferred genetic clusters for *K* =2 and 3.

Table 1. Fixed nucleotide substitutions between *Mesapamea secalis*, *Mesapamea didyma* and *Mesapamea remmi* by genetic markers. Fixed substitutions were observed only in the mitochondrial *COI* gene between *M. secalis* and *M. remmi*, while none of the nuclear genes showed any difference between these two species, hence not enabling an examination of whether *M. remmi* is a hybrid between the two species.

a Five additional sites in *COI* are nearly diagnostic between *M. secalis* versus *M. didyma* and *M. remmi* versus *M. didyma*. COI, cytochrome *c* oxidase subunit I; EF-1a, Elongation factor 1-alpha 1; GAPDH, glyceraldehyde 3 phosphate dehydrogenase; IDH, isocitrate dehydrogenase; MDH, malate dehydrogenase; RpS5, ribosomal protein S5; DDC, Dopa decarboxylase.

Species comparison	*COI* 1446 bp	*EF-1a* 506 bp	*GAPDH* 691 bp	*IDH* 746 bp	*MDH* 407 bp	*RpS5* 626 bp	*wingless* 580 bp	*DDC* 373 bp
M. secalis versus *M. didyma*	4[a]	0	0	0	0	0	0	0
M. secalis versus *M. remmi*	0	0	n/a	n/a	n/a	0	n/a	0
M. didyma versus *M. remmi*	4[a]	0	n/a	n/a	n/a	0	n/a	0

Table 2. *Mesapamea* individuals analysed in this study and a summary of the double digest RAD (ddRAD) sequencing data.

							Loci in assembly		
Species	Sample ID	Total reads ($\times 10^6$)	Reads passed filter ($\times 10^6$)	Clusters at 90%	Retained loci	Mean depth	*m*4	*m*10	*m*17
M. didyma	MM07073	2.84	2.08	25 822	10 648	45.40	5921	3114	334
M. didyma	MM07074	2.05	1.71	22 519	9234	45.55	5349	3025	334
M. didyma	MM07075	2.14	1.70	29 909	11 765	30.16	6322	3173	334
M. didyma	MM07076	3.30	2.72	24 474	9206	48.02	5171	2938	334
M. didyma	MM07077	2.89	2.36	24 075	9208	54.09	4483	2718	334
M. didyma	MM07078	3.37	2.83	23 734	9186	34.8	5457	3034	334
M. remmi	MM07063	2.23	1.84	21 866	7526	45.02	3857	2449	334
M. remmi	MM07064	3.36	2.64	32 371	10 910	55.89	5010	2931	334
M. remmi	MM07065	2.23	1.78	20 136	7782	48.47	4175	2575	334
M. remmi	MM07066	3.55	2.85	31 335	9666	40.27	3584	2335	334
M. remmi	MM07067	5.96	4.95	25 303	11 054	159.93	5713	3137	334
M. remmi	MM07068	2.59	2.09	27 806	11 542	41.35	6168	3206	334
M. secalis	MM07079	2.33	1.95	29 154	11 491	45.71	6433	3236	334
M. secalis	MM07081	0.78	0.61	20 733	4937	22.09	1651	1193	334
M. secalis	MM07082	2.29	1.92	23 707	7965	38.28	4644	2884	334
M. secalis	MM07083	1.96	1.63	19 145	6954	28.13	4071	2671	334
M. secalis	MM07084	1.40	1.17	17 573	6869	22.97	4148	2649	334
Average		**2.66**	**2.17**	**24 686**	**9173**	**47.42**	**4833**	**2780**	**334**

Table 3. Schematic representation of categories of single nucleotide polymorphisms fixed between *Mesapamea didyma* and *Mesapamea remmi* + *Mesapamea secalis.*

	Site				1	2	2	4	5	5	5
					3	1	2	0	5	5	5
		8	9	9	9	3	8	9	0	0	4
		8	0	2	6	8	7	2	1	2	3
Species	Sample ID	0	8	3	2	1	0	3	8	1	7
M. didyma	MM07074	**C**	**T**	**C**	**A**	**T**	**T**	**G**	**T**	**A**	**G**
	MM07075	**C**	**T**	**C**	**A**	**T**	**T**	**G**	**T**	**A**	**G**
	MM07076	**C**	**T**	**C**	**A**	**T**	**T**	**G**	**T**	**A**	**G**
	MM07077	**C**	**T**	**C**	**A**	**T**	**T**	**G**	**T**	**A**	**G**
	MM07078	**C**	**T**	**C**	**A**	**T**	**T**	**G**	**T**	**A**	**G**
	MM07073	**C**	**T**	**C**	**A**	**T**	**T**	**G**	**T**	**A**	**G**
M. remmi	MM07063	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07064	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07065	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07066	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07067	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07068	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
M. secalis	MM07079	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07081	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07082	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07083	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**
	MM07084	**T**	**C**	**T**	**G**	**A**	**C**	**C**	**A**	**C**	**T**

Table 4. Abundance data of *Mesapamea* species in random samples in southern Finland in 2007. Specimens were morphologically identified from the reproductive organs by K-EL.

		Individuals			
Site	Trap type, no. of traps	*M. secalis*	*M. didyma*	*M. remmi*	Total
Hanko	Light 4, bait 6	2507	112	17	2636
		95.1%	4.2%	0.6%	100.0%
Tammisaari	Light 0, bait 1	86	47	1	134
		64.2%	35.1%	0.7%	100.0%
Espoo, Mäkkylä	Light 1, bait 3	99	94	0	193
		51.3%	48.7%	0.0%	100.0%
Espoo, Nuuksio	Light 1, bait 1	43	7	0	50
		86.0%	14.0%	0.0%	100.0%
Virolahti	Light 1, bait 1	159	1	1	161
		98.8%	0.6%	0.6%	100.0%
Joutseno	Light 2, bait 2	299	1	1	301
		99.3%	0.3%	0.3%	100.0%
Total		3193	262	20	3475
		91.9%	7.5%	0.6%	100.0%